CSC446 Homework #1, Kefu Zhu

1. Bishop 1.3

Suppose that we have three coloured boxes r (red), b (blue), and g (green). Box r contains 3 apples, 4 oranges, and 3 limes, box b contains 1 apple, 1 orange, and 0 limes, and box g contains 3 apples, 3 oranges, and 4 limes. If a box is chosen at random with probabilities p(r) = 0.2, p(b) = 0.2, p(g) = 0.6, and a piece of fruit is removed from the box (with equal probability of selecting any of the items in the box).

(A) What is the probability of electing an apple?

$P(Apple) = P(Apple|red) \cdot P(red) + P(Apple|blue) \cdot P(blue) + P(Apple|green) \cdot P(green)$

$= \frac{3}{10} \cdot 0.2 + \frac{1}{2} \cdot 0.2 + \frac{3}{10} \cdot 0.6$

$= 0.3 \cdot 0.2 + 0.5 \cdot 0.2 + 0.3 \cdot 0.6$

$= 0.34$

(B) If we observe that the selected fruit is in fact an orange, what is the probability that it came from the green box?

$P(Orange) = P(Orange|red) \cdot P(red) + P(Orange|blue) \cdot P(blue) + P(Orange|green) \cdot P(green)$

$= \frac{4}{10} \cdot 0.2 + \frac{1}{2} \cdot 0.2 + \frac{3}{10} \cdot 0.6$

$= 0.4 \cdot 0.2 + 0.5 \cdot 0.2 + 0.3 \cdot 0.6$

$= 0.36$

$P(green|Orange) = \frac{P(Orange|green) \cdot P(green)}{P(Orange)} = \frac{0.3 \cdot 0.6}{0.36} = \frac{1}{2}$

2. Bishop 1.11

By setting the derivatives of the log likelihood function (1.54) with respect to μ and σ^2 equal to zero, verify the results (1.55) and (1.56).

(1.54)

$$ln(p(x|\mu, \sigma^2)) = -\frac{1}{2\sigma^2} \sum_{n=1}^{N} (x_n - \mu)^2 - \frac{N}{2} ln\sigma^2 - \frac{N}{2} ln(2\pi)$$

(1.55)

$$\mu_{ML} = \frac{1}{N} \sum_{n=1}^{N} x_n$$

(1.56)

$$\sigma_{ML}^2 = \frac{1}{N} \sum_{n=1}^{N} (x_n - \mu_{ML})^2$$

Proof for 1.55

Calculate the derivative of the log likelihood function (1.54) with respect to μ

$$\frac{\partial}{\partial \mu} ln(p(x|\mu, \sigma^2)) = -\frac{1}{2\sigma} \cdot (-2) \cdot \sum_{n=1}^{N} (x_n - \mu) + 0 + 0 = \frac{1}{\sigma^2} \sum_{n=1}^{N} (x_n - \mu)$$

Maximize the log likelihood function by setting its derivatice to zero

$$\frac{\partial}{\partial \mu} ln(p(x|\mu, \sigma^2)) = \frac{1}{\sigma^2} \sum_{n=1}^{N} (x_n - \mu) = 0$$

$$\Rightarrow \sum_{n=1}^{N} x_n = N \cdot \mu$$

$$\Rightarrow \mu = \frac{1}{N} \sum_{n=1}^{N} x_n$$

Proof for 1.56

Calculate the derivative of the log likelihood function (1.54) with respect to σ

$$\frac{\partial}{\partial \sigma} ln(p(x|\mu, \sigma^2)) = -\frac{1}{2} \cdot (-2) \cdot \sigma^{-3} \sum_{n=1}^{N} (x_n - \mu)^2 - \frac{N}{\sigma} + 0 = \frac{\sum_{n=1}^{N} (x_n - \mu)^2}{\sigma^3} - \frac{N}{\sigma}$$

Maximize the log likelihood function by setting its derivatice to zero

$$\frac{\partial}{\partial \sigma} ln(p(x|\mu, \sigma^2)) = \frac{\sum_{n=1}^{N} (x_n - \mu)^2}{\sigma^3} - \frac{N}{\sigma} = 0$$

$$\Rightarrow \sigma \sum_{n=1}^{N} (x_n - \mu)^2 = N \cdot \sigma^3$$

$$\Rightarrow \sigma^2 = \frac{1}{N} \sum_{n=1}^{N} (x_n - \mu)^2$$

3.

Let (X ⊥⊥ Y) denote that X and Y are independent, and let (X ⊥⊥ Y | Z) denote that X and Y are independent conditioned on Z (Bishop p. 373). Are the following properties true? Prove or disprove.

(A) (X ⫫ W | Z,Y) ∧ (X ⫫ Y | Z) ⇒ (X ⫫ Y,W | Z)

$\because X \perp\!\!\!\perp W \mid Z, Y$

$\therefore P(X \mid Y, W, Z) = P(X \mid Y, Z)$

$\because X \perp\!\!\!\perp Y \mid Z$

$\therefore P(X \mid Y, Z) = P(X \mid Z)$

$\therefore P(X \mid Y, W, Z) = P(X \mid Y, Z) = P(X \mid Z)$

Hence, we proved $X \perp\!\!\!\perp Y, W \mid Z$

(B) (X ⫫ Y | Z) ∧ (X ⫫ Y | W) ⇒ (X ⫫ Y | Z,W)

Let's suppose X, Y, Z are i.i.d. random variables (such as flipping a fair coin) with the following probability

$$\begin{cases} P(X=1) = \frac{1}{2} \\ P(X=-1) = \frac{1}{2} \end{cases}, \begin{cases} P(Y=1) = \frac{1}{2} \\ P(Y=-1) = \frac{1}{2} \end{cases}, \begin{cases} P(Z=1) = \frac{1}{2} \\ P(Z=-1) = \frac{1}{2} \end{cases}$$

In addition, we define event W as $W = XYZ$.

First of all, we need to prove $X \perp\!\!\!\perp Y | Z$ and $X \perp\!\!\!\perp Y | W$ are true in this scenario

- **Proof of $X \perp\!\!\!\perp Y | Z$**

  Since X, Y, Z are i.i.d. random variables, X and Y are independent. Hence we can write $P(X, Y|Z) = P(X|Z)P(Y|Z, X) = P(X|Z)P(Y|Z)$

  $\therefore X \perp\!\!\!\perp Y | Z$ holds

- **Proof of $X \perp\!\!\!\perp Y | W$**

  $P(X = 1 | W = 1) =$

  $P(X = 1, Y = 1, Z = 1 | W = 1) + P(X = 1, Y = 1, Z = -1 | W = 1) +$

  $P(X = 1, Y = -1, Z = 1 | W = 1) + P(X = 1, Y = -1, Z = -1 | W = 1) = \frac{1}{4} + \frac{1}{4} + 0 + 0 = \frac{1}{2}$

  Similarly, we can derive the same result for other combinations of X and W. Same thing for combinations of Y and W. Therefore, we get $P(X|W) = \frac{1}{2}, P(Y|W) = \frac{1}{2}$

  Also, we have

  $P(X = 1, Y = 1 | W = 1) = P(X = 1, Y = 1, Z = 1 | W = 1) + P(X = 1, Y = 1, Z = -1 | W = 1 = \frac{1}{4} + 0 = \frac{1}{4}$

Same logic and computation as above, we eventually can get $P(X, Y|W) = \frac{1}{4}$

$\because P(X, Y|W) = P(X|W)P(Y|W) \ \therefore X \perp\!\!\!\perp Y|W$ holds

Now, let's compute $P(X|Z, W), P(Y|Z, W)$ and $P(X, Y|Z, W)$

$P(X = 1|Z = 1, W = 1) = P(X = 1, Y = 1|Z = 1, W = 1) + P(X = 1, Y = 1|Z = 1, W = 1) = \frac{1}{2} + 0 = \frac{1}{2}$

Similarly, we can derive the same result for other combinations of (X, Z, W). Same thing for combinations of (Y, Z, W). Therefore, we get $P(X|Z, W) = \frac{1}{2}$, $P(Y|Z, W) = \frac{1}{2}$

We can expand $P(X, Y|Z, W)$ as $P(X, Y|Z, W) = P(X|Z, W)P(Y|X, Z, W)$, where $P(X|Z, W) = \frac{1}{2}$ as calculated above, and $P(Y|X, Z, W) = \begin{cases} 1, \ when \ Y = \frac{W}{XZ} \\ 0, \ when \ Y \neq \frac{W}{XZ} \end{cases}$

Therefore, we have

$$P(X, Y|Z, W) = \begin{cases} \frac{1}{2}, \ when \ Y = \frac{W}{XZ} \\ 0, \ when \ Y \neq \frac{W}{XZ} \end{cases}$$

Since $P(X, Y|Z, W)$ does not equal to $P(X|Z, W)P(Y|Z, W) = \frac{1}{4}$, therefore we can conclude that $X \perp\!\!\!\perp Y|Z, W$ does not hold